

HIGHER-LEVEL PATTERN FEATURES PROVIDE ADDITIONAL INFORMATION TO BIRDS WHEN RECOGNIZING AND REJECTING PARASITIC EGGS

Mary Caswell Stoddard^{1*}, Benedict G. Hogan¹, Martin Stevens², Claire N. Spottiswoode^{3, 4}

¹Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544, USA

²Centre for Ecology and Conservation, University of Exeter, Penryn Campus, Penryn, Cornwall, TR10 9FE, UK

³Department of Zoology, University of Cambridge, Cambridge, UK, CB2 3EJ, UK

⁴FitzPatrick Institute of African Ornithology, DST-NRF Centre of Excellence, University of Cape Town, Rondebosch 7701, Cape Town, South Africa

*Author for correspondence: mstoddard@princeton.edu

ABSTRACT

Despite a recent explosion of research on pattern recognition, in both neuroscience and computer vision, we lack a basic understanding of how most animals perceive and respond to patterns in the wild. Avian brood parasites and their hosts provide an ideal study system for investigating the mechanisms of pattern recognition. The cuckoo finch *Anomalospiza imberbis* and its host the tawny-flanked prinia *Prinia subflava* lay highly polymorphic eggs with a great deal of variation in colour and patterning, with the cuckoo finch capable of close egg mimicry. Behavioural experiments in Zambia have previously shown that prinias use colour and multiple “low-level” (occurring in early-stages of visual processing) pattern attributes, derived from spatial frequency analysis, when rejecting foreign eggs. Here we explore the extent to which host birds might also use “higher-level” pattern attributes, derived from a feature detection algorithm, to make rejection decisions. Using a SIFT-based pattern recognition algorithm, *NATUREPATTERNMATCH*, we show that hosts are more likely to reject a foreign egg if its higher-level pattern features – which capture information about the shape and orientation of markings – differ from those of the host eggs. A revised statistical model explains about 37% variance in egg rejection behaviour, and differences in colour, low-level and higher-level pattern features all predict rejection, accounting for 42%, 44% and 14% of the explained variance, respectively. Thus, higher-level pattern features provide a small but measurable improvement to the original model and may be especially useful when colour and low-level pattern features provide hosts with little information. Understanding the relative importance of low- and higher-level pattern features is a valuable goal for future work on animal coloration, especially in the contexts of mimicry, camouflage and individual recognition.

INTRODUCTION

How do animals detect and recognize complex visual patterns, and to what specific features do they respond? Systems in which visual signals have evolved to be deceptive – such as in camouflage and mimicry – provide an especially compelling lens through which to investigate this question, since information is hidden by cheats and retrieved by those often duped [1]. The notorious cheats of the avian world, the brood parasites (together with their hosts) form an ideal system for testing the limits of pattern recognition. For example, visual discrimination by hosts has sometimes resulted in sophisticated colour and pattern mimicry by parasites at the egg, chick, fledgling and even adult stages of the life cycle [2-5]. In response to parasite mimicry, some hosts have evolved more distinctive eggs [6,7] and chicks [8,9], and sometimes better discrimination abilities [10]. Overall, brood parasites and their hosts can be powerful engines of phenotypic diversity, generating an extraordinary range of colourful traits, many of which evolved to thwart host recognition.

For most brood parasite-host pairs, the key evolutionary battle is won or lost at the egg stage [11]. The parasite wins (and host loses) if it successfully sneaks its egg into the host nest, offloading all parental care to the host parents. The host wins (and parasite loses) if it successfully detects and rejects a foreign egg, evading the dual costs of rearing an unrelated chick and (in many cases) losing its own offspring. At this high-stakes moment, what visual cues does a host bird use to detect a parasitic egg? Researchers have vigorously explored this topic in recent years (reviewed in [12]), continuing a long tradition of experimental work on host egg recognition [13,14], but now using visual models appropriate for avian perception. In general, experiments have shown that hosts use avian-perceived differences in egg colour and pattern (spots, markings, speckles) to recognize and reject parasitic eggs [12]. Which specific cues are used, and their relative importance, varies from species to species. Despite these advances, we still have much to learn, especially with respect to pattern.

A decade ago, objective quantification of spatial patterns and texture in studies of animal coloration was rare. Fortunately, this paradigm has changed [15], largely due to the increased use of calibrated digital cameras [16] and accessible image analysis tools [17]. In recent years, several studies on egg recognition have used spatial frequency analysis (sometimes called ‘granularity analysis’) to quantify and compare egg patterns [18,19]. The analysis, which applies a Fourier transform and subsequent filtering to the image, breaks down information into different spatial scales, so that it is easy to quantify basic pattern elements: for example, the average size of the dominant markings (or spots)

and their relative contribution to overall energy (or contrast) in the image. Overall, the process resembles early-stage spatial filtering in vertebrate vision [20] and returns a set of low-level visual features that are probably relevant to an animal's sensory experience [21-23]. But what about more complex features, such as the shapes and orientations of blotches and markings on the eggs? Can birds use these features to identify and reject foreign eggs?

Relatively little is known about spatial vision and pattern recognition in birds [24], in contrast to the extremely well-studied neural mechanisms of birdsong [25]. However, many aspects of vertebrate spatial vision appear to be highly conserved [26], because many animals have neurons whose receptive fields are tuned to different spatial frequencies and orientations [23]. In the early stages of vision, neurons act as spatial filters, breaking down the information in a scene into different spatial frequencies. Next, subsequent neural processes use these filter outputs to identify local features, such as edges and boundaries, which are later combined to form higher-level features like objects. Computationally, these two stages are simulated by spatial frequency analysis (which gives the Fourier power spectrum) and by edge- and feature-detection algorithms, respectively. Granularity analysis stems from the Fourier power spectrum, and there are good reasons for including the low-level metrics it produces in studies of animal colour: they are simple, relevant to animal vision, and easy to quantify [27]. However, without further processing, the granularity spectra alone will not reveal information about edges and objects. For this, edge- and feature detectors are required to build a more complete representation of the scene [27]. Metrics derived from granularity analysis are generally thought to be relevant to low-level vision. Edges and other local features (corners, lines) are also generally considered to be low- or mid-level, while the objects they comprise are higher-level. For simplicity, in this paper we refer to visual features not directly derived from granularity analysis as “higher level,” while acknowledging that this distinction is an oversimplification (see Methods).

Support for the idea that feature detectors might extract visually meaningful information on bird eggs comes from a study by Stoddard and colleagues [6]. Using *NATUREPATTERNMATCH* (NPM), a pattern recognition algorithm based on the Scale-Invariant Feature Transform (SIFT), they showed that host species that are most intensely targeted by common cuckoos *Cuculus canorus* have evolved more recognizable egg pattern signatures. However, this study did not directly relate to egg rejection experiments, and whether pattern features captured by SIFT provide additional information, relative to granularity-based metrics, is unknown. Thus, it remains untested whether birds actually respond to the potential information presented by higher-level pattern features. In addition, no study to date has

tested the relative influence of both low- and higher-level pattern information on egg discrimination behaviour. Indeed, such tests remain rare in any wild system. Egg discrimination behaviour provides an ideal model system for testing the relative importance of low- and higher-level pattern features in a natural setting, since individually distinctive egg markings provide cues with a range of potential information content, and the behavioural response (egg acceptance versus rejection) can be unambiguously scored.

Here, using an Afrotropical brood parasite-host system, we investigate the extent to which host birds use low-level and higher-level pattern features, combined with colour, to identify and reject foreign eggs. We build on a previous study by Spottiswoode and Stevens [19], which used field experiments and avian perceptual modeling to investigate egg rejection behaviour by the tawny-flanked prinia *Prinia subflava*, which is a host commonly parasitized by the cuckoo finch *Anomalospiza imberbis*. This system is characterized by extremely variable eggs in both the parasite and host, differing among individual females in both colour and a variety of pattern markings (figure 1), with the parasite capable of close mimicry. Spottiswoode and Stevens [19] found that egg rejection was predicted by disparity between host and foreign eggs in colour and three low-level pattern parameters (dominant marking size, variability in marking size, dispersion of markings on the egg). Here, using the same dataset, we apply *NATUREPATTERNMATCH* to egg images, allowing us to quantify higher-level pattern features on host and foreign eggs. Combining this with data on colour and low-level pattern features, we build a new model of egg rejection behaviour and evaluate the relative importance of colour, low- and higher-level pattern features to host recognition.

METHODS

Study system and field experiments

Here we briefly review key information about the study system and experimental design; full details can be found in Spottiswoode and Stevens [19]. Egg rejection experiments were performed in January to March 2007–2009 in the Choma District of southern Zambia. In each experiment, a foreign egg was added to the nest of a host prinia female, and one host egg was removed. The foreign experimental egg was laid by a conspecific prinia female, here used as a proxy for a parasitic egg. All eggs were measured with digital calipers and photographed using a Fuji Finepix S7000 camera with a

17% neutral gray card. The colour of one representative egg per host clutch was measured using an Ocean Optics USB2000 spectrophotometer with a PX-2 pulsed xenon light source. Eggs were measured at a 45-degree angle about 5 mm from the probe tip. Experimental clutches were monitored every day or two, and eggs were considered rejected if they disappeared while the rest of the clutch remained intact (*i.e.*, the clutch was not predated), or accepted if they remained in the nest for at least 3 days. Experiments were conducted in $n = 125$ nests; our statistical model is based on 122 nests, after eliminating host clutches for which we had one egg image only, because this did not permit calculation of NPM intraclutch variation. See below and see the Electronic Supplementary Materials (ESM) for details.

Quantifying colour and low-level pattern features

Here we provide a brief overview of the approach used by Spottiswoode and Stevens [19] to quantify egg colour and low-level pattern features. A model of avian (blue tit *Cyanistes caeruleus*) visual discrimination [28] was used to estimate the colour difference (measured in just-noticeable differences, or JNDs) between the host and experimental eggs in each experiment. JNDs less than 1.00 suggest that two colours cannot be discriminated; larger values suggest that two colours should be more easily discriminated. Next, egg images were rescaled to 50 pixels/mm using egg measurements taken in the field in order to enable meaningful comparisons based on marking size/texture scale. Egg images were linearized (so that pixel values scale uniformly with light captured by the camera's sensor) and equalized (so that the camera's channels respond equally to a gray standard, allowing the image to be converted to reflectance) [16]. Grayscale images from the green channel were produced because the mediumwave sensor corresponds relatively closely to an avian luminance channel, which is thought to be important for achromatic/texture perception [29].

Using a modified approach from Stoddard and Stevens [18], granularity analysis was performed on each egg image. Regions-of-interest from the narrow, middle and wide regions of the egg were selected and then subjected to fast Fourier transform and band-pass filtering, resulting in a granularity spectrum for each egg region. The spectrum (figure 1b) illustrates how much information (or overall pattern "energy") in the image is present at different spatial scales, with small filter sizes corresponding to large, low-spatial frequency markings, and large filter sizes corresponding to small, high-spatial frequency markings. From the granularity spectrum, metrics describing the egg's pattern

can be calculated. The spatial scale with the largest contribution to overall pattern energy corresponds to the dominant marking size (“pattern filter size”), and the proportion of the total energy contained at this spatial scale (“pattern proportion energy”) provides a measure of how dominant the marking size is. An additional measure not based on the granularity spectrum was calculated: the degree to which markings are clustered toward the wide end of the egg (“pattern dispersion”). This was calculated by thresholding each egg image (egg markings = 1; ground colour = 0) and quantifying the difference in the proportion of markings between the two poles of the egg, by subtracting the proportion value of the narrow region from that of the wide region. We consider this to be a low-level pattern measure.

Luminance (achromatic) contrast, several additional pattern metrics, and morphometric and life history metrics (*e.g.*, egg length, egg breadth, clutch size, incubation stage) were also quantified by Spottiswoode and Stevens [19]; however, they were not significant predictors of egg rejection, and so for clarity we do not discuss them further.

Quantifying higher-level pattern features using NATUREPATTERNMATCH

To extract features using *NATUREPATTERNMATCH* (NPM), we applied histogram equalization and median filtering to egg images. Histogram equalization is a method for contrast enhancement in images, allowing areas of low local contrast to gain higher contrast. Median filtering reduces noise in images while preserving edges. These modifications allowed us to compensate for less-than-perfect image quality and lighting by enhancing otherwise “washed out” pattern features.

Full details of the NPM program can be found in [6]. In brief, the program uses the Scale-Invariant Feature Transform (SIFT; [30,31]) to detect local features in an image. On egg patterns, SIFT features are associated with egg maculation (blotches and markings) and provide information about a marking’s shape, contrast and dominant orientation. Each SIFT feature (pink arrows in figure 1b) is represented visually by a vector whose magnitude corresponds to the feature’s dominant scale (or size) and direction corresponds to the feature’s dominant orientation. The vector shows the feature’s dominant size and orientation only: in reality each feature encodes information about the local pixels in a 128-dimensional feature vector (figure 1b), capturing small details about a marking’s shape and contrast. After extracting features, NPM then matches populations of features from one egg image to another, resulting in a similarity score for each pair of images. Importantly, NPM’s image-to-image

pattern matching process is inspired by texture similarity rather than object identity: the idea is to compare two similar (but not identical) natural patterns.

SIFT is an algorithm for detecting, describing and matching low- and mid-level visual features (*i.e.*, points, edges, blobs, corners, patches). To detect features, the image is first convolved with filters that resemble the circularly symmetrical Difference-of-Gaussian receptive fields of retinal neurons [31]. This step approximates the Laplacian of Gaussian operation, which is used in classic edge detection approaches [32] and is helpful because it can be used to identify stable image features. The next steps of the algorithm were not specifically designed to mimic biological vision, but were inspired by the response properties of complex neurons in the visual cortex: these neurons retain their orientation and spatial frequency specificity even if a feature changes position slightly [30]. The algorithm works to identify visual features that are generally of intermediate complexity (*i.e.*, more complex than a simple bar or line but less complex than a face) and share properties with features to which cortical neurons respond and use for object recognition (in primates) [30]. SIFT features are generally invariant (to a degree) to changes in size, rotation or location, so that a given feature can be matched to another feature (on a different egg) even if it appears somewhat warped, rescaled or rotated. These are the kinds of features that might be important for recognition.

Individual SIFT features are local (low- or mid-level) features. In many applications, collections of SIFT features are used to identify higher-level objects, but NPM uses them in a more generic way to assess texture similarity (or difference). We consider suites of SIFT features as used by NPM, which capture aspects of egg texture, to be “higher-level” pattern descriptors, in order to distinguish them from the low-level features derived from spatial filtering/granularity analysis (*i.e.*, energy in different spatial frequency bands) and from individual SIFT features. We therefore consider measures of the similarity or difference between suites of SIFT features (as assessed by NPM, which compares both the constituent features themselves and the number of features – see below) to be higher-level measures of egg pattern similarity or difference.

We quantified patterns using NPM on real host eggs, some of which were used as experimental “parasite” eggs ($n = 375$ eggs in 122 nests). We used the pairwise dissimilarity values generated by NPM (1-similarity) to calculate the following:

1. NPM experimental-host distance: the average pairwise pattern distance as calculated by NPM between the experimental egg and each of the eggs in the host clutch (figure 1c). Here, the host clutch includes eggs in the nest at the time of the experiment, *after* removal of one of the host eggs (see ***Study system and field experiments*** above).

To investigate any predictive effect of the number of SIFT features extracted by NPM on experimental and host eggs, we also calculated the following measure:

2. NPM experimental-host difference in feature number: the mean absolute difference in the number of features detected on the experimental egg and the number of features detected on the eggs of the host clutch. Again, the host clutch includes eggs in the nest at the time of the experiment, *after* removal of one of the host eggs.

Measures generated by NPM are based on pairwise distances (differences) between two eggs and provide a measure of difference. Therefore, “NPM experimental-host distance” and “NPM experimental-host difference in feature number” are calculated as the average pairwise distance between the experimental egg and all eggs in the host clutch. It is not possible, for example, to calculate average NPM pattern features on an “idealized” host egg by averaging across multiple eggs. In Spottiswoode & Stevens [19], low-level pattern metrics across eggs in a clutch were calculated slightly differently: these measures (pattern filter size, pattern proportion energy, pattern dispersion) were averaged across host eggs to calculate – for a given host clutch – the average pattern filter size, pattern proportion energy, and pattern dispersion, which were then compared to the experimental egg to yield a clutch-level experimental-host distance. However, the way we calculate experimental-host distance across eggs in a clutch in the present study provides a comparable measure of the average similarity between the host eggs and the foreign egg.

Our NPM-based pattern measures (above) are based on host eggs after the experimental removal of one host egg, consistent with the approach used by Spottiswoode and Stevens [19], which was motivated by the observation that cuckoo finches remove one or more host eggs from the clutch when they lay their own. However, we also calculated the NPM-based measures (above) using *all eggs in the host clutch* (including the one that was experimentally removed), consistent with the idea that prinia hosts use a template-based mechanism (at least in part) for recognizing their own eggs and rejecting foreign eggs [33]. Calculated this way, a host would have access to higher-level pattern

features on *all of the eggs it laid*, since this total information is relevant to the host's innate or learned template. These metrics were similar regardless of the method of calculation (Pearson's $r=0.898$, $CI_{0.95}=0.857$ to 0.928 , $t=22.328$, $df = 120$, $p<.001$) and yielded qualitatively similar results for predicting rejection (see ESM table 5). For simplicity, in the main text we discuss only the NPM-based pattern measures described above, calculated *not including* the removed host egg.

To confirm that the pattern difference between *real* cuckoo finch eggs and host eggs ("NPM parasite-host distance") is approximated by the experimental equivalent, which is based on prinia eggs only, we calculated an additional measure: "NPM parasite-host distance." This measure is a simulated estimate of the distance between a parasite egg and host eggs, accomplished by finding the distance between each host clutch and a randomly selected cuckoo finch egg ($n=85$). While this simulation used some cuckoo finch eggs in multiple comparisons, it estimates the distribution of distances possible and is indeed comparable to NPM experimental-host distance (figure 1c). Similarly, Spottiswoode and Stevens [19] demonstrated conspecific prinia eggs are suitable surrogates for real parasitic eggs. Our simulation approach assumes that cuckoo finches lay their eggs haphazardly in host nests (*i.e.*, they do not lay eggs in host clutches that provide the best phenotype match to their own), consistent with field observations of cuckoo finches [19].

In addition, whereas Spottiswoode and Stevens [19] exclusively examined differences between host and foreign eggs, here we also measured host intraclutch and interclutch variation, as measured by NPM (see ESM tables 1-2). This allowed us to describe the degree of higher-level pattern variation in the host population. It also allowed us to test whether host rejection behaviour is influenced by the host's egg pattern signature and/or by the pattern difference between the host's eggs and a foreign egg. In theory, a host with low intraclutch variation and/or high interclutch variation would have a strong egg pattern signature [6] which, independent of the parasite's egg, could contribute to successful egg recognition and rejection. We calculated these measures as follows:

3. NPM intraclutch distance: the average pairwise distance between a given host egg and all other host eggs in the clutch, averaged to the clutch level. This measure includes all host eggs laid by the host female, and not the experimental egg (figure 1c). This is a measure of how repeatable an individual female host's own egg patterns are.
4. NPM interclutch distance: the average pairwise distance between a given host egg in clutch 1, and all host eggs in each clutch $[2...N]$, averaged across clutches $[2....N]$, which is then

averaged for all eggs in clutch 1 giving a clutch-level measure. N is the total number of clutches in the population. This measure includes all host eggs laid by the host female, and not the experimental egg (figure 1c). This is a measure of how distinctive an individual female host's egg patterns are relative to other females in the population.

Statistical analyses

To generate a predictive model of egg rejection behaviour, we used logistic regression (function `glm`) in R [34]. We included the following parameters in the initial model: NPM experimental-host distance, NPM experimental-host difference in feature number, pattern proportion energy, pattern dispersion, pattern filter size, and colour. In all analyses, all predictors were standardized (to have mean zero and standard deviation one). This had no effect on the results but facilitated interpretation of regression coefficients. The untransformed predictors are shown in figures.

We then used a stepwise model selection technique, on the basis of the Akaike information criterion, to drop the least important terms from a comprehensive model containing all predictors until the best model was selected (function `stepAIC` in R package `MASS`). Following Spottiswoode and Stevens [19], the explanatory power of the parameters retained in the final model was assessed using hierarchical partitioning (package `hier.part` in R), and the total explanatory power of the model was calculated using Nagelkerke's R squared using the package `rsq` in R [35,36].

In exploratory analysis presented in the supplementary information, we found only limited support for interactions between model terms (ESM table 3), so for clarity only main effects were considered in the main text. We observed little correlation between predictors that were retained in the final model (ESM table 4). In the initial comprehensive model, there was a strong correlation between NPM experimental-host distance and NPM experimental-host difference in feature number (Pearson's $r=0.641$, $CI_{0.95}=0.523$ to 0.735 , $t=9.143$, $df=120$, $p<.001$), which is discussed below.

Finally, we isolated the effects of colour, low and higher-level pattern features to investigate their respective contributions to the final model. A given predictor's influence on the model's output (predicted probability of rejection; using the `predict.glm` function in R package `glm`) can be eliminated by setting the input values for that predictor to its mean (because all predictors are standardized). This allowed us to use the same model and coefficients to predict egg rejection in the (simulated)

absence of one or more predictors. To contrast the predictive power of models including and excluding higher- and lower-level pattern measures, we generated two datasets and used these to predict rejection. In one dataset, the data were unchanged from those in the selected final model except that NPM experimental-host distance was fixed at its mean value (here rejection was predicted only on the basis of colour and low-level pattern measures). In the other dataset, colour and low-level pattern measures were fixed to their respective mean values (here rejection was predicted only on the basis of NPM experimental-host distance). A similar approach was used to produce plots of the conditional effect of each predictor in turn (figure 2a-e).

RESULTS

Quantifying egg pattern differences using NPM

Using NPM to compare eggs within and among clutches (figure 1c & 1d), we found that intraclutch variation was lower on average than interclutch variation, although it was distributed over a much greater range. Differences in the two distributions indicated that clutches were on average more similar to themselves than to other clutches in terms of pattern. The distribution of NPM experimental-host distances was intermediate. In addition, there was overlap between NPM experimental-host distances and NPM simulated parasite-host distances; however, cuckoo finch egg patterns were slightly more different on average to host clutches than the experimental (conspecific prinia) eggs.

What visual features are used by hosts to recognize and reject eggs?

A binomial generalized linear model predicting egg rejection with each of the final predictors from Spottiswoode and Stevens [19] – colour, pattern dispersion, pattern proportion energy, and pattern filter size – successfully replicated the previously published findings [19].

Model selection on the basis of AIC indicated that the best model of egg rejection contained only the previous predictors and NPM experimental-host distance (*i.e.*, NPM experimental-host difference in number was not retained; see table 1, figure 2, see also ESM). This model explained 36.7% of the variance in egg rejection, compared to 31.9% reported in Spottiswoode and Stevens [19]. Including NPM experimental-host distance in the model resulted in a small but measurable improvement to the original model. Much of the variance in egg rejection behaviour (~63%) remains unexplained. In the

revised model, colour difference remained the most important predictor, making up 42% of the explained variance, followed by pattern proportion energy (16%), pattern dispersion (17%), NPM experimental-host distance (14%) and pattern filter (marking) size (11%).

For completeness, we also produced a model using NPM experimental-host distance only; it was a significant predictor of egg rejection ($df=120$, $z=2.447$, $p=.014$), explaining ~7% of the variance in egg rejection. In addition, we produced a model with NPM experimental-host distance and colour only: both were significant predictors of egg rejection, as expected ($df=119$, NPM experimental-host distance, $z=2.654$, $p=.007$; Colour, $z=3.353$, $p=.001$), explaining ~20% of the variance. Therefore, SIFT-based features can predict rejection behaviour in the absence of colour and/or low-level pattern metrics, but the overall model is better when colour, low-level pattern and higher-level pattern measures are all included. Critically, the significant parameters predicting egg rejection (table 1) are not highly correlated with one another (ESM table 4; note one significant weak correlation between NPM experimental-host distance and dispersion, Pearson's $r=0.276$, $CI_{0.95}=0.1$ to 0.432 , $t=3.143$, $df=120$, $p<.01$), which suggests that NPM is capturing non-redundant pattern information. In other words, the information derived from NPM (about distinctive features with shapes and orientations) is different from that derived from spatial frequency analysis (dominant marking size and its relative contribution to the overall pattern contrast).

In our initial analyses, we found that NPM experimental-host distance and NPM experimental-host difference in feature number were highly correlated (see Methods). When we excluded NPM experimental-host distance from the initial model, NPM experimental-host difference in feature number was not retained by stepwise model selection; by contrast, when we excluded NPM experimental-host difference in feature number from the model, NPM experimental-host distance was retained in the final model. This suggests that NPM experimental-host distance, while correlated with the absolute difference in feature number between eggs, contains information about the shapes, gradients and orientations of features that is useful for recognition. Just comparing the number of features on host eggs to the number of features on the experimental egg does not significantly predict egg rejection.

In the main text, we concentrated on NPM experimental-host distance (and NPM experimental-host difference in feature number) because it is most directly comparable to the measures used in Spottiswoode and Stevens [19] in that it compares the appearance of foreign and host eggs.

Additional analyses that included host intraclutch and interclutch variation as predictors of egg rejection can be found in the ESM (tables 1-2). Neither significantly predicted egg rejection.

When are higher-level features most important?

Figure 4 demonstrates how colour, low-level and higher-level pattern features predict rejection. Red dots (representing rejected eggs) in Quadrant 1 indicate cases in which NPM experimental-host distance does predict rejection (defined as predicting a probability that the egg will be rejected >50%, based on NPM experimental-host distance when the other predictor variables are set to their mean values), but colour and low-level pattern metrics do not. These rejected eggs are represented by red lines in figure 2f-j, which suggest that when pattern proportion energy, pattern dispersion, and pattern filter size provide little information (small differences between host and experimental eggs), large differences in higher-level pattern features (as estimated by NPM experimental-host distance) might contribute to rejection. One specific example of this is shown in clutch 1 of figure 3 (see Discussion). Small differences in NPM experimental-host distance do not usually impede rejection if there is sufficient information in the other channels (Quadrant 4 in figure 4; clutches 2-3 in figure 3). See Discussion for additional details.

DISCUSSION

Understanding how animals perceive and prioritize lower- and higher-level pattern information is a pressing goal for sensory ecologists [27,37], with few explicit tests of this idea in wild animals. Tawny-flanked prinias, a frequent host of the parasitic cuckoo finch, have been shown previously to use multiple visual cues when discriminating against foreign eggs [19], making them ideal subjects for an in-depth analysis of pattern recognition. In this study, we demonstrate that both low-level and higher-level pattern information – together with colour – predict egg rejection behaviour in prinias. We find that including a measure of visual difference (between experimental and host eggs) estimated using higher-level pattern information in a model with colour and low-level pattern features increases the explained variance in rejection by about 5%, from 32% to 37%. This result suggests that pattern metrics derived from collections of SIFT features, which capture details about the shape and orientation of egg markings, provide novel, non-redundant information to birds. Overall, using a mix of low- and higher-level pattern information may improve the chances of successful egg recognition and rejection. The increase in predictive power in the model when adding higher-level information is

relatively small, and only 14% of the explained variance in rejection is based on differences in higher-level features between host and experimental eggs. This suggests that egg rejection behaviour is primarily driven by colour and low-level pattern differences. Higher-level features nonetheless provide some additional information and appear to be as important for rejection as any single low-level pattern measure (each accounting for about 11–17% of explained variance in rejection).

When are higher-level pattern features likely to be particularly beneficial? In our egg rejection experiments, there were many trials in which prinias rejected experimental eggs even though differences in low-level pattern metrics were small (figure 3, clutch 1) and colour and low-level metrics together were poor predictors of egg rejection (figure 4, rejected eggs in Quadrants 1 and 3). In many of these cases, rejected eggs were substantially different from host eggs in terms of SIFT features, such that NPM experimental-host distance was high or moderate (figure 4, rejected eggs in Quadrant 1 and the top of Quadrant 3). However, there were some cases in which high NPM experimental-host distance appears to have been insufficient to trigger rejection in the absence of large differences in colour and low-level pattern metrics (figure 4, accepted eggs in Quadrant 1; see also additional analyses and discussion in figure ESM 1). In addition, some eggs were rejected when colour, low-level and higher-level feature differences were *all* poor predictors of rejection (figure 4, rejected eggs in Quadrant 3; see also additional analyses and discussion in figure ESM 2). What might explain these results? Unexpected egg acceptance might be attributed to host parents that were naïve first-time breeders that had not yet acquired a reliable template of their own egg's appearance. Unexpected egg rejection might be due to hosts' ability to make full use of other cues of parasitism such as the sight of adult brood parasites in the environment, as prinias are known to do [38].

Taken together, these observations suggest that when colour and low-level pattern metrics provide little information to a host, differences in higher-level pattern features may *sometimes* be uniquely informative for identifying a parasite egg. One such scenario is highlighted in clutch 1 in figure 3, showing eggs from a trial in which the experimental egg was rejected. The colour and low-level pattern differences between the experimental and host eggs are all low; the dominant markings on the eggs are of similar size, and the patterning (though much denser on host eggs) is distributed relatively evenly across the narrow and wide egg regions. By contrast, the dissimilarity of experimental and host eggs when measured with NPM is high, indicating that higher-level features likely contributed to successful egg rejection.

These observations raise the questions of whether and how host birds might prioritize certain types of visual information. It is as yet unknown whether there is some sort of hierarchy of features used in rejection, such that hosts first use colour, then low-level pattern information, and then – if those are uninformative – they switch to higher-level features. Alternatively, hosts may somehow integrate all features simultaneously, but weight some cues as more important than others (like colour). Our predictive model shows that colour and low-level pattern metrics account for 42% and 44%, respectively, of the overall explained variance in rejection rate (Table 1), while higher-level pattern features account for 14%. One possibility is that colour and low-level pattern features are evaluated first, consistent with the idea that initial colour and low-level pattern processing occurs in the very early stages of vision [39]. If these two channels do not provide large visual differences, higher-level pattern features – processed slightly downstream – might be assessed and sometimes (but not always) tip the balance in favor of egg rejection (figure 4, Quadrant 1). In other words, when the forgery is very good, higher-level pattern assessment might be necessary to spot the fake. For a fascinating parallel in the art world, consult [40], in which low-level pattern metrics (spatial frequency analysis with wavelets) were sufficient for spotting a notorious van Gogh forgery, but a more complex method (using multidimensional histograms, much like SIFT features) was required to successfully identify other non-van Gogh paintings. So do birds process visual information in a sequential manner, using higher-level pattern features only if required? In all likelihood, vision in birds – as in humans [41,42] – is not strictly feedforward, but rather characterized by rich interactions and feedbacks. Therefore, colour, low-level and higher-level pattern information are probably integrated in complex ways.

We find that the significant parameters predicting egg rejection (table 1) are not highly correlated with one another (ESM table 4), indicating that colour, low-level and higher-level pattern features may have evolved to maximize information content on host eggs, as suggested previously [7,18,19]. Moreover, this finding underscores the fact that NPM is capturing non-redundant pattern information (*i.e.*, different from the information represented by low-level features derived from spatial frequency analysis or measures of pattern dispersion). Interestingly, we did not find that egg rejection behaviour was predicted by a host's intraclutch or interclutch variation, as measured by NPM (ESM tables 1-2). This may be because prinia hosts appear to base their rejection decisions on the pattern cues that differ most reliably between real and parasitic eggs in the population [19], so that hosts are tuned to the traits that will most consistently reveal a large difference between parasite and host eggs irrespective of the host's own egg phenotype. In other words, whether or not a host has a good egg

pattern signature (low intraclutch variation and/or high interclutch variation) matters less than the host's ability to compare effectively its own eggs to the parasite egg. Prinias know the appearance of their own eggs, rejecting foreign eggs when they differ from an internal template [33]. Puzzlingly, prinias apparently fail to use the one 100% reliable pattern cue – scribbled lines (figure 1a) that they always produce (in highly variable quantities among females) but cuckoo finches do not – to reject parasitic eggs [19], because unscribbled parasite eggs are routinely accepted. However, prinias might detect and use scribbled “information” more than we realize. The SIFT algorithm readily finds features along scribbled lines (figure 3), and some differences between a scribbled and unscribbled egg would be captured by NPM experimental-host distance. Accordingly, the presence (or absence) of scribbled lines might affect rejection in subtle ways that we do not yet appreciate.

Rapid advances in image analysis and computer vision have catapulted us into a new era of animal coloration research: the age of pattern quantification, growing steadily for years, has fully arrived. Pattern analysis methods, ranging from the simple (spatial filtering) to the complex (deep neural networks), are proliferating, and it will be crucial for biologists to consider the biological relevance of such methods [27]. Ultimately, detailed behavioural experiments on birds and other animals will be needed to test which techniques are the most appropriate and informative. Where possible, these should include both low- and higher-level pattern features. A recent study of camouflage [37], for example, showed that low- and higher-level pattern metrics together predicted a human's ability to detect an artificial prey item on a screen, but this depended on the type of camouflage. Discovering whether camouflage and mimicry evolved in response to similar (or different) forms of low- and higher-level pattern detection is an exciting prospect for the future. Because brood parasites and their hosts have evolved ever-more elaborate visual tricks to outwit the other, they form an ideal model system for this area of inquiry.

TABLES

Table 1 – Predictors of egg rejection in experimentally parasitized nests. I% is the proportion of overall variance explained by the model (36.7%) contributed by each variable independently. VIF indicates variance inflation factors for each predictor. Degrees of freedom = 116.

	Estimate	SE	95% CI		Z	P	I(%)	VIF
(Intercept)	0.110	0.220	-0.317	0.553	0.500	0.617		
NPM experimental-host distance	0.504	0.238	0.052	0.992	2.118	0.034*	14.313	1.089
Pattern proportion energy	0.605	0.251	0.146	1.137	2.411	0.016*	15.586	1.053
Pattern dispersion	0.570	0.236	0.121	1.057	2.411	0.016*	17.496	1.099
Pattern filter size	0.480	0.224	0.047	0.933	2.142	0.032*	10.830	1.043
Color	1.064	0.281	0.555	1.663	3.782	0.0002**	41.776	1.163

Sig. Levels: ** $p < .01$, * $p < .05$

FIGURES

Figure 1. A) Eggs laid by different prinia females (outer circle) and cuckoo finch females (inner circle). Photo by CNS, previously published in [43]. B) Illustration of feature extraction by NPM (using SIFT) and granularity analysis alongside colour measurements. Pattern dispersion is not derived from granularity analysis and is calculated separately. C) Diagram illustrating the methods for calculating pairwise measures. From left to right: Host intraclutch distance; here calculated as the average of the blue lines. Host interclutch distance; here the blue lines are averaged to find the distance between the left clutch and the top-right clutch. Following this, the red lines are averaged to find the distance between the left clutch and bottom-right clutch. The mean of these two measurements is the interclutch distance of the left-hand clutch. Experimental-host distance; here the blue lines are averaged to find the distance between the host clutch (right) and experimental egg (left, red). Simulated parasite-host distance; here the blue lines are averaged to find the distance between the host clutch (right) and a cuckoo finch egg (left, green; see main text). D) Boxplot of the overall distribution of distance metrics based on NPM extraction of SIFT features: NPM intraclutch and NPM interclutch distances, NPM experimental-host distance, and NPM simulated parasite-host distance.

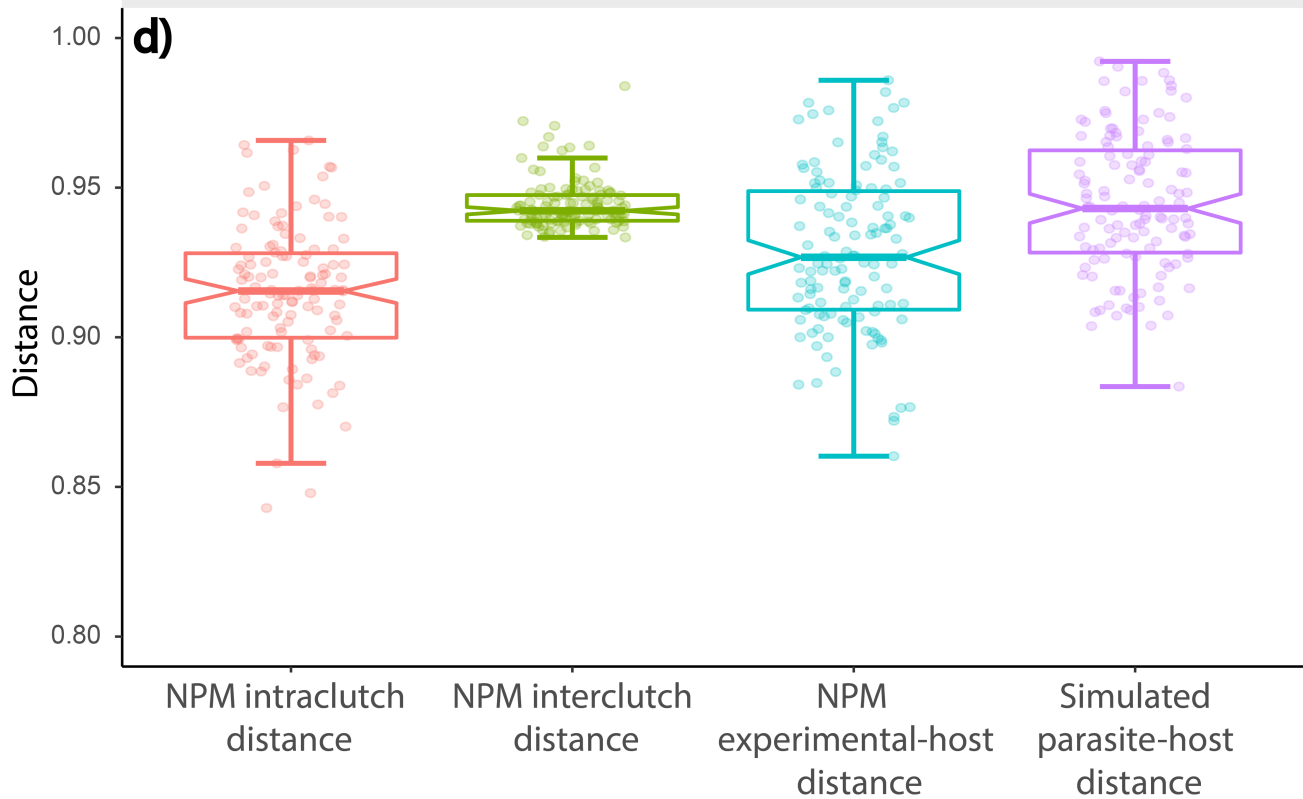
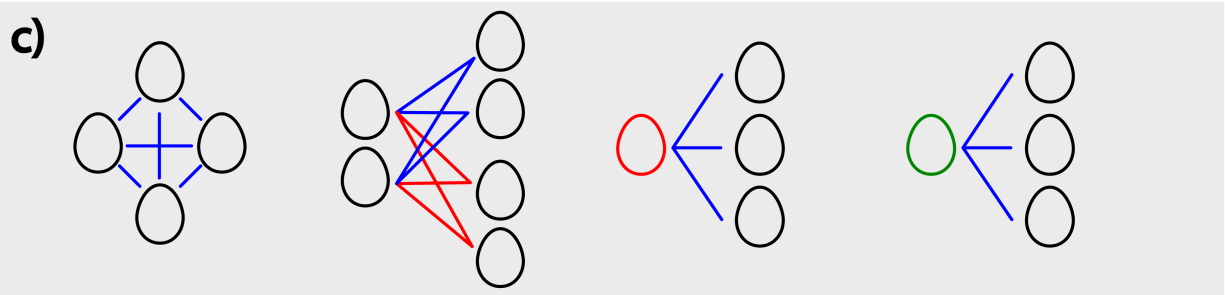
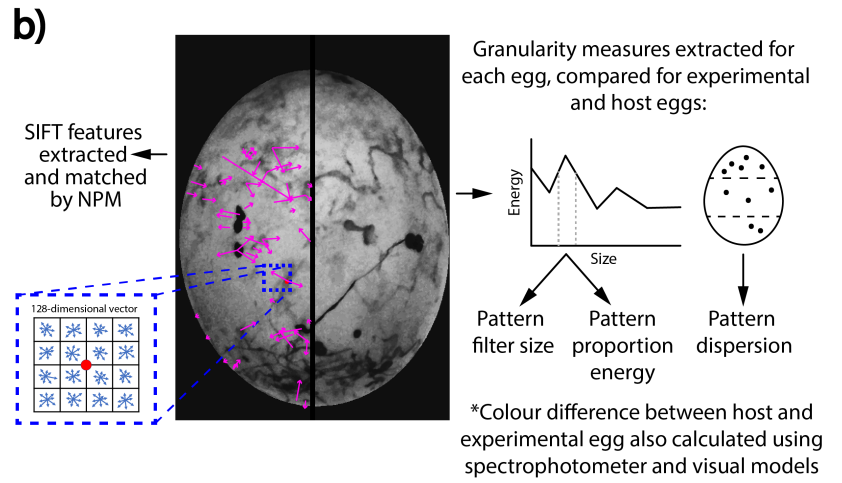
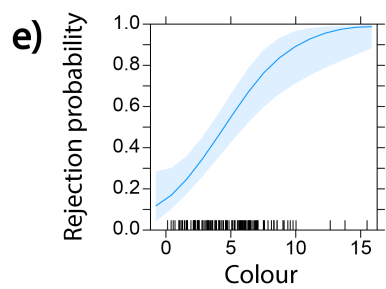
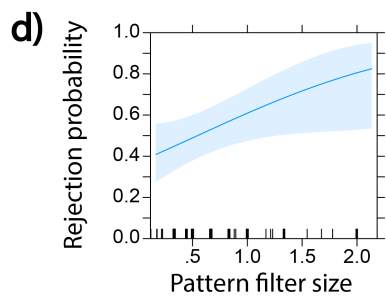
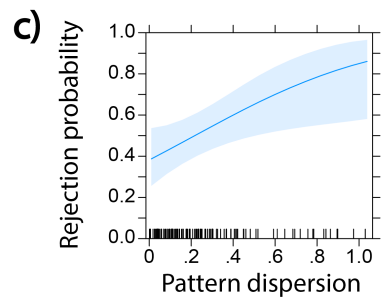
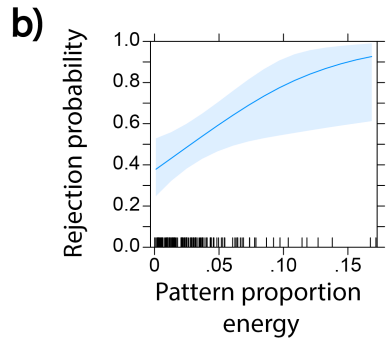
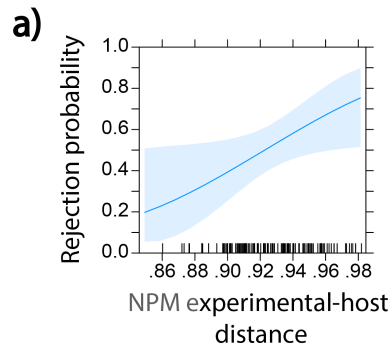


Figure 2. a-e) Plots of the conditional effect of each predictor in the selected model. Y-axis is predicted probability of rejection, x-axis is predictor value. Plotted effects indicate the effect of changes in predictor value when all other predictors are held at their mean values (see main text). Shaded areas represent 95% confidence intervals. Tick marks on the x-axis indicate the distribution of predictor values. For all predictors, the predicted probability of rejection increases with greater measured difference between experimental and host eggs. f-j) Normalized histograms of the values of each predictor retained in the final model; blue vertical line indicates the mean value. Red vertical lines indicate values for rejected eggs that fall in Quadrant 1 of figure 4; these refer to eggs for which NPM experimental-host distance predicts rejection, but low-level pattern and colour differences do not.

Conditional Effects



Data Distribution

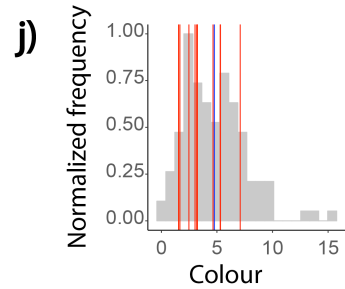
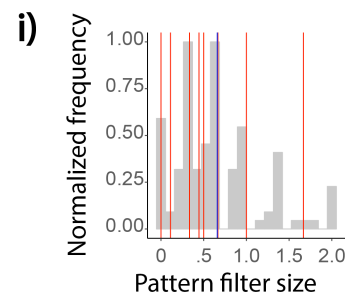
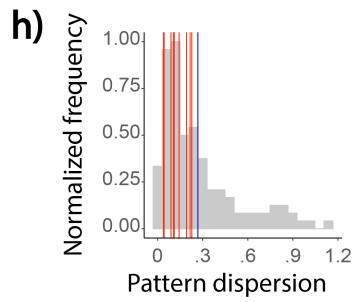
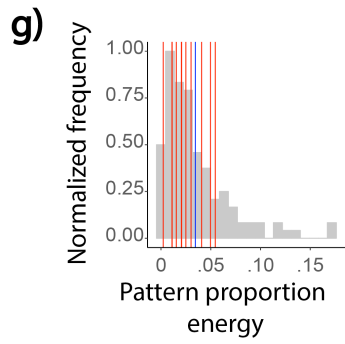
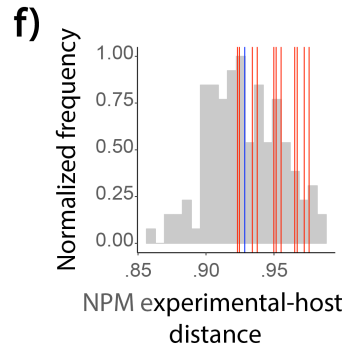


Figure 3. Selected example host/experimental egg combinations. Each row of images represents one egg rejection trial, with experimental eggs on the right. Also noted are the values of higher-level (NPM) pattern, low-level pattern and colour measures of difference between experimental and host eggs. Three selected trials illustrate cases where 1) NPM experimental-host distance is high but low-level differences are low; 2) low-level differences are high, but NPM experimental-host distance is low; 3) all pattern differences are low, but colour difference between experimental and host eggs is high. The experimental egg was rejected in all three of these experiments.

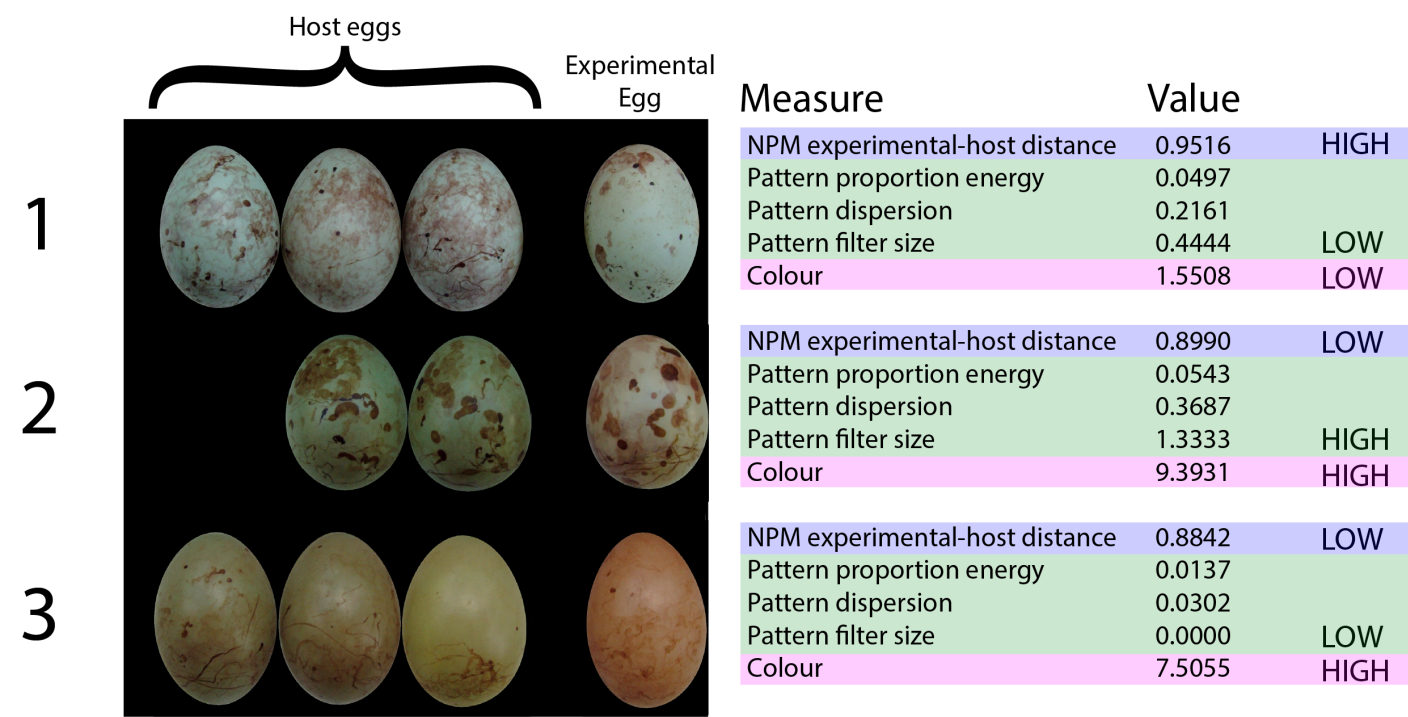
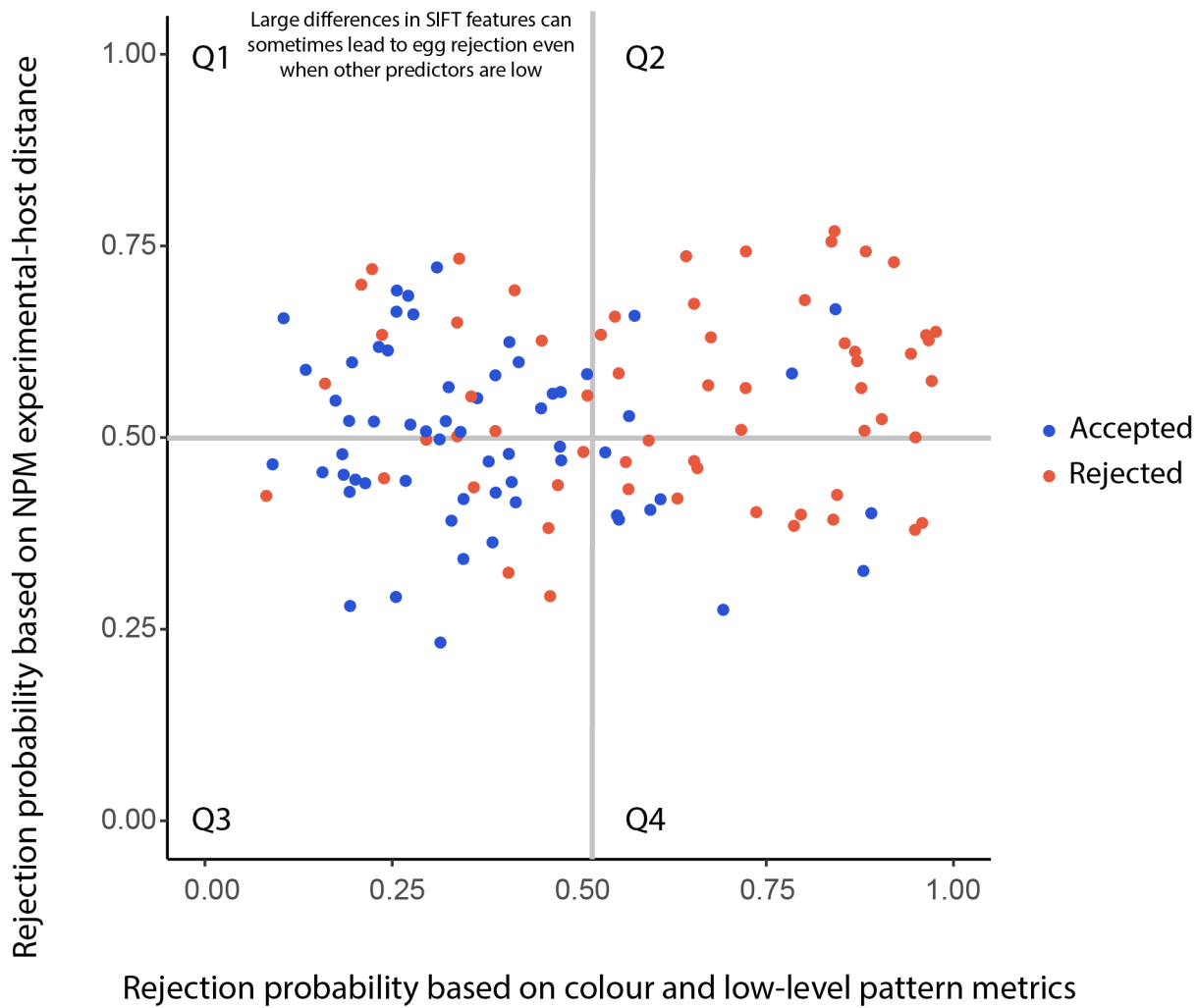


Figure 4. Predicted egg rejection on the basis of colour and low-level pattern metrics only (x-axis, other predictors fixed to mean) or NPM experimental-host distance (y-axis, other predictors fixed to mean).



ACKNOWLEDGEMENTS: We thank the Editors, two anonymous reviewers, Daniel Hanley and Tanmay Dixit for helpful feedback on the manuscript. In the field, we thank the many people who found prinia nests and who allowed access to their land, and the Zambia Wildlife Authority for permits. MCS was supported by Princeton University and a Sloan Research Fellowship. CNS was supported by a BBSRC David Phillips Fellowship (BB/J014109/1), and fieldwork was supported by the DST-NRF Centre of Excellence at the FitzPatrick Institute, a Royal Society Dorothy Hodgkin Research Fellowship and Sidney Sussex College Junior Research Fellowship to CNS.

AUTHOR CONTRIBUTIONS: MCS and CNS planned the study. BGH and MCS directed the NPM and statistical analyses, which were conducted by BGH. MCS and BGH wrote the manuscript, to which all authors contributed. CNS conducted fieldwork. CNS and MS collected and analyzed the original data.

REFERENCES

1. Stevens M. 2016 Cheats and deceptions: how animals and plants exploit and mislead. Oxford University Press.
2. Langmore NE, Spottiswoode CN. 2012 Visual trickery in avian brood parasites. In *Host Manipulation by Parasite* (eds DP Hughes, J Brodeur, F Thomas), pp. 95–118. Oxford, UK: Oxford University Press.
3. Davies NB. 2011 Cuckoo adaptations: trickery and tuning. *J. Zool.* **284**, 1–14. (doi:10.1111/j.1469-7998.2011.00810.x)
4. Stoddard MC, Hauber ME. 2017 Colour, vision and coevolution in avian brood parasitism. *Phil. Trans. R. Soc. B* **372**, 20160339. (doi:10.1098/rstb.2016.0339)
5. Feeney WE, Troschianko J, Langmore NE, Spottiswoode CN. 2015 Evidence for aggressive mimicry in an adult brood parasitic bird, and generalized defences in its host. *Proc. R. Soc. London Ser. B.* **282**, 20150795. (doi:10.1111/j.2007.0908-8857.03813.x)
6. Stoddard MC, Kilner RM, Town C. 2014 Pattern recognition algorithm reveals how birds evolve individual egg pattern signatures. *Nat. Comm.* **5**, 4117. (doi:10.1038/ncomms5117)
7. Caves EM, Stevens M, Iversen ES, Spottiswoode CN. 2015 Hosts of avian brood parasites have evolved egg signatures with elevated information content. *Proc. R. Soc. London Ser. B.* **282**, 20150598. (doi:10.1098/rstb.1995.0073)
8. Sato NJ, Tanaka KD, Okahisa Y, Yamamichi M, Kuehn R, Gula R, Ueda K, Theuerkauf J. 2015 Nestling polymorphism in a cuckoo-host system. *Curr. Biol.* **25**, R1164–R1165. (doi:10.1016/j.cub.2015.11.028)
9. Attisano A, Sato NJ, Tanaka KD, Okahisa Y, Kuehn R, Gula R, Ueda K, Theuerkauf J. 2018 Visual discrimination of polymorphic nestlings in a cuckoo-host system. *Sci. Rep.*, **8**:10359 (doi:10.1038/s41598-018-28710-5)
10. Spottiswoode CN, Stevens M. 2011 How to evade a coevolving brood parasite: egg discrimination versus egg variability as host defences. *Proc. Biol. Sci.* **278**, 3566–3573.

(doi:10.1098/rspb.2011.0401)

11. Kilner RM. 2006 The evolution of egg colour and patterning in birds. *Biol. Rev.* **81**, 383. (doi:10.1017/S1464793106007044)
12. Honza M, Cherry MI. 2017 Egg characteristics affecting egg rejection. In *Avian Brood Parasitism* (eds M Soler), pp. 401–419. Springer, Cham. (doi:10.1007/978-3-319-73138-4_22)
13. Davies N, Brooke ML. 1988 Cuckoos versus reed warblers: adaptations and counteradaptations. *Anim. Behav.* **36**, 262–284.
14. Rothstein SI. 1982 Mechanisms of avian egg recognition: Which egg parameters elicit responses by rejecter species? *Behav. Ecol. Sociobiol.* **11**, 229–239. (doi:10.1007/BF00299299)
15. Cuthill IC, Allen WL, Arbuckle K, Caspers B, Chaplin G, Hauber ME, Hill GE, Jablonski NG, Jiggins CD, Kelber A, et al. 2017 The biology of color. *Science* **357**, 6350. (doi:10.1126/science.aan0221)
16. Stevens M, Párraga CA, Cuthill IC, Partridge JC, Troscianko TS. 2007 Using digital photography to study animal coloration. *Biol. J. Linn. Soc.* **90**, 211–237. (10.1111/j.1095-8312.2007.00725.x)
17. Troscianko J, Stevens M. 2015 Image calibration and analysis toolbox - a free software suite for objectively measuring reflectance, colour and pattern. *Methods Ecol Evol*, **6**, 1320–1331. (doi:10.1111/2041-210X.12439)
18. Stoddard MC, Stevens M. 2010 Pattern mimicry of host eggs by the common cuckoo, as seen through a bird's eye. *Proc. R. Soc. London Ser. B.* **277**, 1387–1393. (doi:10.1098/rspb.2009.2018)
19. Spottiswoode CN, Stevens M. 2010 Visual modeling shows that avian host parents use multiple visual cues in rejecting parasitic eggs. *Proc. Natl. Acad. Sci. USA* **107**, 8672–8676. (doi:10.1073/pnas.0910486107)
20. Campbell FW, Robson JG. 1968 Application of fourier analysis to the visibility of gratings. *J. Physiol. Lond.* **197**, 551–566. (doi:10.1113/jphysiol.1968.sp008574)
21. Godfrey D, Lythgoe, JN, Rumball DA. 1987 Zebra stripes and tiger stripes: the spatial frequency distribution of the pattern compared to that of the background is significant in display and crypsis. *Biol. J. Linn. Soc.* **32**, 427–433. (doi:10.1111/j.1095-8312.1987.tb00442.x)
22. Pérez-Rodríguez L, Jovani R, Stevens M. 2017 Shape matters: animal colour patterns as signals of individual quality. *Proc. R. Soc. London Ser. B.* **284**, 20162446–10. (doi:10.1098/rspb.2016.2446)
23. Stevens M. 2007 Predator perception and the interrelation between different forms of protective coloration. *Proc. R. Soc. London Ser. B.* **274**, 1457–1464. (doi:10.1098/rspb.2007.0220)

24. Soto FA, Wasserman EA. 2011 Visual object categorization in birds and primates: Integrating behavioral, neurobiological, and computational evidence within a 'general process' framework. *Cogn. Affect. Behav. Neurosci.* **12**, 220–240. (doi:10.3758/s13415-011-0070-x)
25. Bolhuis JJ, Gahr M. 2006 Neural mechanisms of birdsong memory. *Nat. Rev. Neurosci.* **7**, 347–357.
26. Cronin, TW, Johnsen, S, Marshall, NJ, Warrant, EJ. 2014 *Visual ecology*. Princeton University Press.
27. Stoddard MC, Osorio D. *In Press*. Animal coloration patterns: linking spatial vision to quantitative analysis. *American Naturalist*.
28. Vorobyev M, Osorio D. 1998 Receptor noise as a determinant of colour thresholds. *Proc. R. Soc. London Ser. B.* **265**, 351-358.
29. Jones CD, Osorio D. 2004 Discrimination of oriented visual textures by poultry chicks. *Vision Res.* **44**, 83–89. (doi:10.1016/j.visres.2003.08.014)
30. Lowe D. 2000 Towards a computational model for object recognition in IT cortex. In *International Workshop on Biologically Motivated Computer Vision*, pp. 20–31. Berlin, Germany: Springer.
31. Lowe DG. 2004 Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision* **60**, 91–110.
32. Marr D, Hildreth E. 1980 Theory of edge detection. *Proc. R. Soc. London Ser. B.* **207**, 187–217.
33. Stevens M, Troscianko J, Spottiswoode, CN. 2013 Repeated targeting of the same hosts by a brood parasite compromises host egg rejection. *Nat. Comm.* **4**, 2475. (doi:10.1038/ncomms3475)
34. R Development Core Team 2015 2015 R: A language and environment for statistical computing.
35. Faraway JJ. 2016 Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models. Chapman & Hall/CRC Press: Boca Raton, FL.
36. Nagelkerke NJ. 1991 A note on a general definition of the coefficient of determination. *Biometrika* **78**, 691–692.
37. Troscianko J, Skelhorn J, Stevens M. 2017 Quantifying camouflage: how to predict detectability from appearance. *BMC Evol. Biol.* **17**, 1–13. (doi:10.1186/s12862-016-0854-2)
38. Feeney WE, Troscianko J, Langmore NE, Spottiswoode CN. 2015. Evidence for aggressive mimicry in an adult brood parasitic bird, and generalised defences in its host. *Proc. R. Soc. London Ser. B.* **282**: 2015079.

39. Snowden R, Snowden RJ, Thompson P, Troscianko T. 2012 *Basic vision: an introduction to visual perception*. Oxford University Press.
40. Johnson C, Hendriks E, Bereznoy I, Brevdo E, Hughes S, Daubechies I, Li J, Postma E, Wang J. 2008 Image processing for artist identification. *IEEE Signal Process. Mag.* **25**, 37–48. (doi:10.1109/MSP.2008.923513)
41. Cloutman LL. 2013 Interaction between dorsal and ventral processing streams: where, when and how? *Brain and Language* **127**, 251–263. (doi:10.1016/j.bandl.2012.08.003)
42. Medathati NVK, Neumann H, Masson GS, Kornprobst P. 2016 Bio-inspired computer vision. *Comput. Vis. Image Und.* **150**, 1–30. (doi:10.1016/j.cviu.2016.04.009)
43. Spottiswoode CN, Stevens M. 2012. Host-parasite arms races and rapid changes in bird egg appearance. *Am. Nat.* **1279**, 633-648. (doi: 10.1086/665031).